

# Anomalous recurrence properties of many-dimensional zero-drift random walks

Nicholas Georgiou<sup>\*†</sup>   Mikhail V. Menshikov<sup>\*</sup>   Aleksandar Mijatović<sup>‡</sup>  
 Andrew R. Wade<sup>\*</sup>

29 June 2015

## Abstract

Famously, a  $d$ -dimensional, spatially homogeneous random walk whose increments are non-degenerate, have finite second moments, and have zero mean is recurrent if  $d \in \{1, 2\}$  but transient if  $d \geq 3$ . Once spatial homogeneity is relaxed, this is no longer true. We study a family of zero-drift spatially non-homogeneous random walks (Markov processes) whose increment covariance matrix is asymptotically constant along rays from the origin, and which, in any ambient dimension  $d \geq 2$ , can be adjusted so that the walk is either transient or recurrent. Natural examples are provided by random walks whose increments are supported on ellipsoids that are symmetric about the ray from the origin through the walk's current position; these *elliptic random walks* generalize the classical homogeneous Pearson–Rayleigh walk (the spherical case). Our proof of the recurrence classification is based on fundamental work of Lamperti.

*Key words:* Non-homogeneous random walk; elliptic random walk; zero drift; recurrence; transience.

*AMS Subject Classification:* 60J05 (Primary) 60J10, 60G42, 60G50 (Secondary)

## 1 Introduction

A  $d$ -dimensional random walk that proceeds via a sequence of unit-length steps, each in an independent and uniformly random direction, is sometimes called a *Pearson–Rayleigh* random walk (PRRW), after the exchange in the letters pages of *Nature* between Karl Pearson and Lord Rayleigh in 1905 [16]. Pearson was interested in two dimensions and questions of migration of species (such as mosquitoes) [17], although Carazza has speculated that Pearson was a golfer [3, p. 419]; Rayleigh had earlier considered the acoustic ‘random walks’ in phase space produced by combinations of sound waves of the same amplitude and random phases.

---

<sup>\*</sup>Department of Mathematical Sciences, Durham University, South Road, Durham DH1 3LE, UK.

<sup>†</sup>Heilbronn Institute for Mathematical Research, School of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW.

<sup>‡</sup>Department of Mathematics, Imperial College London, 180 Queen’s Gate, London SW7 2AZ, UK.

The PRRW can be represented via partial sums of sequences of i.i.d. random vectors that are uniformly distributed on the unit sphere  $\mathbb{S}^{d-1}$  in  $\mathbb{R}^d$ . Clearly the increments have mean zero, i.e., the PRRW has *zero drift*. The PRRW has received some renewed interest recently as a model for microbe locomotion [1, 14, 15]. Chapter 2 of [8] gives a general discussion of these walks, which have been well-understood for many years. In particular, it is well known that the PRRW is recurrent for  $d \in \{1, 2\}$  and transient if  $d \geq 3$ .

Suppose that we replace the spherically symmetric increments of the PRRW by increments that instead have some *elliptical* structure, while retaining the zero drift. For example, one could take the increments to be uniformly distributed on the surface of an ellipsoid of fixed shape and orientation, as represented by the picture on the right of Figure 1. More generally, one should view the ellipses in Figure 1 as representing the *covariance* structure of the increments of the walk (we will give a concrete example later; the uniform distribution on the ellipse is actually not the most convenient for calculations).

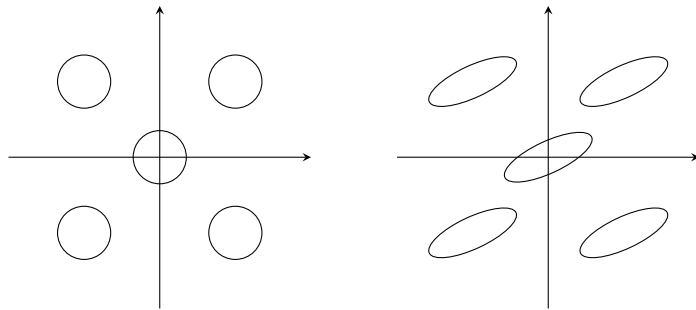


Figure 1: Pictorial representation of spatially homogeneous random walks with increments distributed on a fixed circle (*left*) and a fixed ellipse (*right*).

A little thought shows that the walk represented by the picture on the right of Figure 1 is essentially no different to the PRRW: an affine transformation of  $\mathbb{R}^d$  will map the walk back to a walk whose increments have the same covariance structure as the PRRW. To obtain genuinely different behaviour, it is necessary to abandon spatial homogeneity.

In this paper we consider a family of spatially *non-homogeneous* random walks with zero drift. These include generalizations of the PRRW in which the increments are not i.i.d. but have a distribution supported on an ellipsoid of fixed size and shape but whose orientation depends upon the current position of the walk. Figure 2 gives representations of two important types of example, in which the ellipsoid is aligned so that its principal axes are parallel or perpendicular to the vector of the current position of the walk, which sits at the centre of the ellipse.

The random walks represented by Figure 2 are no longer sums of i.i.d. variables. These modified walks can behave very differently to the PRRW. For instance, one of the two-dimensional random walks represented in Figure 2 is *transient* while the other (as in the classical case) is recurrent. The reader who has not seen this kind of example before may take a moment to identify which is which. It is this *anomalous recurrence behaviour* that is the main subject of the present paper. In the next section, we give a formal description of our model and state our main results.

We end this introduction with a brief comment on motivation. In biology, the PRRW is more natural than a lattice-based walk for modelling the motion of microscopic organ-

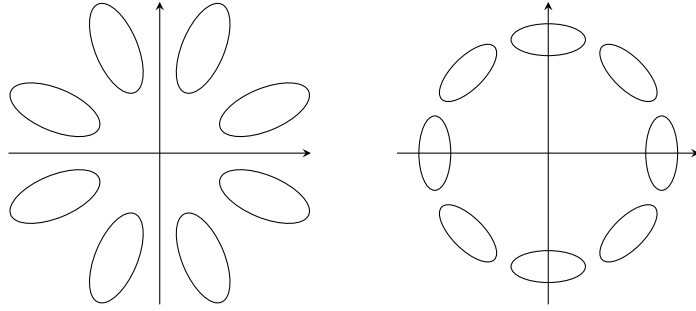


Figure 2: Pictorial representation of spatially non-homogeneous random walks with increments distributed on a radially-aligned ellipse with major axis aligned in the radial sense (*left*) and in the transverse sense (*right*).

isms, such as certain bacteria, on a surface. Experiment suggests that the locomotion of several kinds of cells consists of roughly straight line segments linked by discrete changes in direction: see, e.g., [14, 15]. The generalization to elliptically-distributed increments studied here represents movement on a surface on which either radial or transverse motion is inhibited. In chemistry and physics, the trajectory of a finite-step PRRW (also called a ‘random chain’) is an idealized model of the growth of weakly interacting polymer molecules: see, e.g., §2.6 of [8]. The modification to ellipsoid-supported jumps represents polymer growth in a biased medium.

## 2 Model and main results

We work in  $\mathbb{R}^d$ ,  $d \geq 1$ . Our main interest is in  $d \geq 2$ , as we shall explain shortly. Write  $\mathbf{e}_1, \dots, \mathbf{e}_d$  for the standard orthonormal basis vectors in  $\mathbb{R}^d$ . Write  $\mathbf{0}$  for the origin in  $\mathbb{R}^d$ , and let  $\|\cdot\|$  denote the Euclidean norm and  $\langle \cdot, \cdot \rangle$  the Euclidean inner product on  $\mathbb{R}^d$ . Write  $\mathbb{S}^{d-1} := \{\mathbf{u} \in \mathbb{R}^d : \|\mathbf{u}\| = 1\}$  for the unit sphere in  $\mathbb{R}^d$ . For  $\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ , set  $\hat{\mathbf{x}} := \mathbf{x}/\|\mathbf{x}\|$ ; also set  $\hat{\mathbf{0}} := \mathbf{e}_1$ , for convenience. For definiteness, vectors  $\mathbf{x} \in \mathbb{R}^d$  are viewed as column vectors throughout.

We now define  $X = (X_n, n \in \mathbb{Z}_+)$ , a discrete-time, time-homogeneous Markov process on a (non-empty, unbounded) subset  $\mathbb{X}$  of  $\mathbb{R}^d$ . Formally,  $(\mathbb{X}, \mathcal{B}_{\mathbb{X}})$  is a measurable space,  $\mathbb{X}$  is a Borel subset of  $\mathbb{R}^d$ , and  $\mathcal{B}_{\mathbb{X}}$  is the  $\sigma$ -algebra of all  $B \cap \mathbb{X}$  for  $B$  a Borel set in  $\mathbb{R}^d$ . Suppose  $X_0$  is some fixed (i.e., non-random) point in  $\mathbb{X}$ . Write

$$\Delta_n := X_{n+1} - X_n \quad (n \in \mathbb{Z}_+)$$

for the increments of  $X$ . By assumption, given  $X_0, \dots, X_n$ , the law of  $\Delta_n$  depends only on  $X_n$  (and not on  $n$ ); so often we ease notation by taking  $n = 0$  and writing just  $\Delta$  for  $\Delta_0$ . We also use the shorthand  $\mathbb{P}_{\mathbf{x}}[\cdot] = \mathbb{P}[\cdot \mid X_0 = \mathbf{x}]$  for probabilities when the walk is started from  $\mathbf{x} \in \mathbb{X}$ ; similarly we use  $\mathbb{E}_{\mathbf{x}}$  for the corresponding expectations.

We make the following moments assumption:

(A0) There exists  $p > 2$  such that  $\sup_{\mathbf{x} \in \mathbb{X}} \mathbb{E}_{\mathbf{x}}[\|\Delta\|^p] < \infty$ .

The assumption (A0) ensures that  $\Delta$  has a well-defined mean vector  $\mu(\mathbf{x}) := \mathbb{E}_{\mathbf{x}}[\Delta]$ , and we suppose that the random walk has *zero drift*:

(A1) Suppose that  $\mu(\mathbf{x}) = \mathbf{0}$  for all  $\mathbf{x} \in \mathbb{X}$ .

The assumption (A0) also ensures that  $\Delta$  has a well-defined covariance matrix, which we denote by  $M(\mathbf{x}) := \mathbb{E}_{\mathbf{x}}[\Delta\Delta^\top]$ , where  $\Delta$  is viewed as a column vector. To rule out pathological cases, we assume that  $\Delta$  is *uniformly non-degenerate*, in the following sense.

(A2) There exists  $v > 0$  such that  $\text{tr}(M(\mathbf{x})) = \mathbb{E}_{\mathbf{x}}[\|\Delta\|^2] \geq v$  for all  $\mathbf{x} \in \mathbb{X}$ .

Note that assumption (A2) is weaker than *uniform ellipticity*, which in this context usually means, for some  $\varepsilon > 0$ ,  $\mathbb{P}_{\mathbf{x}}[\Delta \cdot \mathbf{u} \geq \varepsilon] \geq \varepsilon$  for all  $\mathbf{u} \in \mathbb{S}^{d-1}$  and all  $\mathbf{x}$ .

Our main interest is in a recurrence classification. First, we state the following basic ‘non-confinement’ result.

**Proposition 2.1.** *Suppose that  $X$  satisfies assumptions (A0), (A1) and (A2). Then*

$$\limsup_{n \rightarrow \infty} \|X_n\| = +\infty, \text{ a.s.} \quad (2.1)$$

We give the proof of Proposition 2.1 in Section 4; we actually prove more, namely that the hypotheses of Proposition 2.1 ensure that a version of Kolmogorov’s ‘other’ inequality holds. The fact (2.1) ensures that questions of the escape of trajectories to infinity are non-trivial. Indeed, we will give conditions under which one or other of the following two behaviours (which are not *a priori* exhaustive) occurs:

- $\lim_{n \rightarrow \infty} \|X_n\| = +\infty$ , a.s., in which case we say that  $X$  is *transient*;
- $\liminf_{n \rightarrow \infty} \|X_n\| \leq r_0$ , a.s., for some constant  $r_0 \in \mathbb{R}_+$ , when we say  $X$  is *recurrent*.

If  $X$  is an irreducible time-homogeneous Markov chain on a locally finite state-space, these definitions reduce to the usual notions of transience and recurrence; in general state-spaces, our approach allows us to avoid unnecessary technicalities concerning irreducibility.

In dimension  $d = 1$ , it is a consequence of the classical Chung–Fuchs theorem (see [4] or Chapter 9 of [9]) that a spatially *homogeneous* random walk with zero drift is necessarily recurrent. However, this is *not* true for a spatially non-homogeneous random walk: as observed by Rogozin and Foss [19], a counterexample is provided by a version of the ‘oscillating random walk’ of Kemperman [10] in which the increment law is one of two distributions (with mean zero but infinite second moment) depending on the walk’s present sign. Our conditions exclude these heavy-tailed phenomena, so that in  $d = 1$  recurrence is assured in our setting.

**Theorem 2.2.** *Suppose that  $d = 1$ . Suppose that  $X$  satisfies assumptions (A0), (A1), and (A2). Then  $X$  is recurrent.*

Theorem 2.2 is essentially contained in a result of Lamperti [12, Theorem 3.2]; we give a self-contained proof below. Theorem 2.2 shows that in  $d = 1$ , under mild conditions, the classical Chung–Fuchs recurrence classification for homogeneous zero-drift random walks extends to zero-drift non-homogeneous random walks. The purpose of the present paper is to demonstrate a natural family of examples in dimension  $d \geq 2$  where this extension fails, and hence exhibit the following.

**Fact.** *There exist spatially non-homogeneous random walks whose increments are non-degenerate, have uniformly bounded second moments, and have zero mean, which are*

- *transient in  $d = 2$ ;*
- *recurrent in  $d \geq 3$ .*

Although certainly appreciated by experts, this fact is perhaps not as widely known as it might be. Zeitouni (pp. 91–92 of [20]) describes an example of a transient zero-drift random walk on  $\mathbb{Z}^2$ , and states that the idea “goes back to Krylov (in the context of diffusions)”. Peres, Popov and Sousi [18] investigate the minimal number of different increment distributions required for anomalous recurrence behaviour.

We now introduce our family of non-homogeneous random walks. Write  $\|\cdot\|_{\text{op}}$  for the matrix (operator) norm given by  $\|M\|_{\text{op}} = \sup_{\mathbf{u} \in \mathbb{S}^{d-1}} \|M\mathbf{u}\|$ . The following assumption on the asymptotic stability of the covariance structure of the process along rays is central.

**(A3)** Suppose that there exists a positive-definite matrix function  $\sigma^2$  with domain  $\mathbb{S}^{d-1}$  such that, as  $r \rightarrow \infty$ ,

$$\varepsilon(r) := \sup_{\mathbf{x} \in \mathbb{X}: \|\mathbf{x}\| \geq r} \|M(\mathbf{x}) - \sigma^2(\hat{\mathbf{x}})\|_{\text{op}} \rightarrow 0.$$

A little informally, (A3) says that  $M(\mathbf{x}) \rightarrow \sigma^2(\hat{\mathbf{x}})$  as  $\|\mathbf{x}\| \rightarrow \infty$ ; in what follows, we will often make similar statements, formal versions of which may be cast as in (A3).

Note that (A2) and (A3) together imply that  $\text{tr}(\sigma^2(\mathbf{u})) \geq v > 0$ ; next we impose a key assumption on the form of  $\sigma^2$  that is considerably stronger. To describe this, it is convenient to introduce the notation  $\langle \cdot, \cdot \rangle_{\mathbf{u}}$  that defines, for each  $\mathbf{u} \in \mathbb{S}^{d-1}$ , an inner product on  $\mathbb{R}^d$  via

$$\langle \mathbf{y}, \mathbf{z} \rangle_{\mathbf{u}} := \mathbf{y}^\top \cdot \sigma^2(\mathbf{u}) \cdot \mathbf{z} = \langle \mathbf{y}, \sigma^2(\mathbf{u}) \cdot \mathbf{z} \rangle, \quad \text{for } \mathbf{y}, \mathbf{z} \in \mathbb{R}^d.$$

**(A4)** Suppose that there exist constants  $U$  and  $V$  with  $0 < U < V < \infty$  such that, for all  $\mathbf{u} \in \mathbb{S}^{d-1}$ ,

$$\langle \mathbf{u}, \mathbf{u} \rangle_{\mathbf{u}} = U, \quad \text{and} \quad \text{tr}(\sigma^2(\mathbf{u})) = V.$$

Informally,  $V$  quantifies the total variance of the increments, while  $U$  quantifies the variance in the radial direction; necessarily  $U \leq V$ . The assumption that  $0 \neq U \neq V$  excludes some degenerate cases. As we will see, one possible way to satisfy condition (A4) is to suppose that the eigenvectors of  $\sigma^2(\mathbf{u})$  are all parallel or perpendicular to the vector  $\mathbf{u}$ , and that the corresponding eigenvalues are all constant as  $\mathbf{u}$  varies; the level sets of the corresponding quadratic forms  $q_{\mathbf{u}}(\mathbf{x}) := \langle \mathbf{x}, \mathbf{x} \rangle_{\mathbf{u}}$  for  $\mathbf{u} \in \mathbb{S}^{d-1}$  are then ellipsoids like those depicted in Figure 2.

Our main result is the following, which shows that both transience and recurrence are possible for *any*  $d \geq 2$ , depending on parameter choices; as seen in Theorem 2.2, this possibility of anomalous recurrence behaviour is a genuinely multidimensional phenomenon under our regularity conditions.

**Theorem 2.3.** *Suppose that  $X$  satisfies (A0)–(A4), with constants  $0 < U < V$  as defined in (A4). The following recurrence classification is valid.*

- (i) *If  $2U < V$ , then  $X$  is transient.*
- (ii) *If  $2U > V$ , then  $X$  is recurrent.*
- (iii) *If  $2U = V$  and (A3) holds with  $\varepsilon(r) = O(r^{-\delta_0})$  for some  $\delta_0 > 0$ , then  $X$  is recurrent.*

Moreover, we show that in any of the above cases,  $X$  is *null* in the following sense.

**Theorem 2.4.** *Suppose that  $X$  satisfies (A0)–(A4), with constants  $0 < U < V$  as defined in (A4). Then, in any of the cases (i)–(iii) in Theorem 2.3, for any bounded  $A \subset \mathbb{R}^d$ ,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k \in A\} = 0, \text{ a.s. and in } L^q \text{ for any } q \geq 1. \quad (2.2)$$

*Remark 2.5.* Theorems 2.3 and 2.4 both remain valid if we permit  $V = U > 0$  in (A4); indeed, the condition  $U < V$  is not used in the proof of Theorem 2.3 given below, so this case is recurrent, by Theorem 2.3(ii). The condition  $U < V$  is used at one point to simplify the proof of Theorem 2.4 given below, but a small modification of the argument also works in the case  $U = V$ .

The remainder of the paper is organised as follows. In Section 3 we describe a specific family of examples called *elliptic random walk models* that satisfy assumptions (A0)–(A4) and exhibit both transient and recurrent behaviour dependent on the parameters of the model. We also present some simulated data that depicts the random walks in both cases. In Section 4 we prove a  $d$ -dimensional martingale version of Kolmogorov’s other inequality and use that to prove the non-confinement result (Proposition 2.1). In Section 5 we prove the recurrence classification (Theorem 2.3), and in Section 6 we prove Theorem 2.4. In the appendix we prove recurrence in the one-dimensional case (Theorem 2.2).

Finally, we remark that in work in progress we investigate diffusive scaling limits for random walks of the type described in the present paper; the diffusions that appear as scaling limits possess certain pathologies from the point of view of diffusion theory that make them interesting in their own right.

### 3 Example: Elliptic random walk model

Let  $d \geq 2$ . We describe a specific model on  $\mathbb{X} = \mathbb{R}^d$  where the jump distribution at  $\mathbf{x} \in \mathbb{R}^d$  is supported on an ellipsoid having one distinguished axis aligned with the vector  $\mathbf{x}$ . The model is specified by two constants  $a, b > 0$ . Construct  $\Delta$  as follows. Given  $X_0 = \mathbf{x}$ , take  $\zeta$  uniform on  $\mathbb{S}^{d-1}$  and set

$$\Delta = Q_{\hat{\mathbf{x}}} D \zeta \quad (3.1)$$

for  $Q_{\hat{\mathbf{x}}}$  an orthogonal matrix representing a transformation of  $\mathbb{R}^d$  mapping  $\mathbf{e}_1$  to  $\hat{\mathbf{x}}$ , and  $D = \sqrt{d} \operatorname{diag}(a, b, \dots, b)$ . See Figure 3.

(Recall that  $\hat{\mathbf{0}} = \mathbf{e}_1$ , so for  $X_0 = \mathbf{0}$  we can take  $Q_{\hat{\mathbf{x}}} = I$  and  $\Delta = D\zeta$ .) Thus  $\Delta$  is a random point on an ellipsoid that has one distinguished semi-axis, of length  $a\sqrt{d}$ , aligned in the  $\hat{\mathbf{x}}$  direction, and all other semi-axes of length  $b\sqrt{d}$ . Note that the law of  $\Delta$  is well defined owing to the spherical symmetry of the uniform distribution on  $\mathbb{S}^{d-1}$  and the fact that only one axis of the ellipsoid is distinguished (for this reason it is enough to take any  $Q_{\hat{\mathbf{x}}}$  satisfying  $Q_{\hat{\mathbf{x}}} \mathbf{e}_1 = \hat{\mathbf{x}}$  in order to define  $\Delta$ ; see also Remark 3.1 below).

Note also that  $\Delta$  is not chosen to be uniformly distributed on the surface of the ellipsoid; this does not affect the range of asymptotic behaviour exhibited by the family of walks as  $a$  and  $b$  vary, but it does simplify the calculation of  $M(\mathbf{x})$ . Indeed, we have

$$M(\mathbf{x}) = \mathbb{E}_{\mathbf{x}}[\Delta \Delta^\top] = \mathbb{E}[Q_{\hat{\mathbf{x}}} D \zeta \zeta^\top D Q_{\hat{\mathbf{x}}}^\top] = Q_{\hat{\mathbf{x}}} D \mathbb{E}[\zeta \zeta^\top] D Q_{\hat{\mathbf{x}}}^\top = \frac{1}{d} Q_{\hat{\mathbf{x}}} D^2 Q_{\hat{\mathbf{x}}}^\top,$$

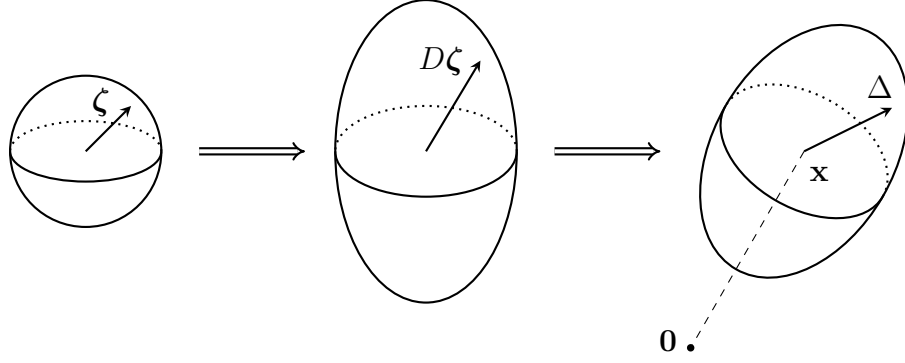


Figure 3: Definition of  $\Delta = Q_{\hat{\mathbf{x}}} D \zeta$ .

by linearity of expectation, and using the fact that  $\mathbb{E}[\zeta \zeta^\top] = \frac{1}{d} I$  for  $\zeta$  uniformly distributed on  $\mathbb{S}^{d-1}$ . Also, a calculation similar to the above confirms that  $\mu(\mathbf{x}) = \mathbf{0}$  for all  $\mathbf{x} \in \mathbb{R}^d$ , since  $\mathbb{E}[\zeta] = \mathbf{0}$ .

Since  $\|\Delta\|$  is bounded above by  $\sqrt{d} \max\{a, b\}$ , assumption (A0) holds. Clearly (A1) and (A3) hold, with  $\sigma^2(\mathbf{u}) = \frac{1}{d} Q_{\mathbf{u}} D^2 Q_{\mathbf{u}}^\top$  for  $\mathbf{u} \in \mathbb{S}^{d-1}$ . It is also a simple matter to check that (A2) and (A4) hold: the matrix  $\sigma^2(\mathbf{u})$  represented in coordinates for the orthonormal basis  $\{Q_{\mathbf{u}} \mathbf{e}_1 = \mathbf{u}, Q_{\mathbf{u}} \mathbf{e}_2, \dots, Q_{\mathbf{u}} \mathbf{e}_d\}$  is diagonal with entries  $a^2, b^2, \dots, b^2$ . Indeed,

$$\begin{aligned} \sigma^2(\mathbf{u}) &= \frac{1}{d} Q_{\mathbf{u}} D^2 Q_{\mathbf{u}}^\top = Q_{\mathbf{u}} [b^2 I + (a^2 - b^2) \mathbf{e}_1 \mathbf{e}_1^\top] Q_{\mathbf{u}}^\top \\ &= a^2 \mathbf{u} \mathbf{u}^\top + b^2 (I - \mathbf{u} \mathbf{u}^\top), \end{aligned}$$

and therefore  $\langle \mathbf{u}, \mathbf{u} \rangle_{\mathbf{u}} = \langle \mathbf{u}, \sigma^2(\mathbf{u}) \cdot \mathbf{u} \rangle = a^2 > 0$  for all  $\mathbf{u} \in \mathbb{S}^{d-1}$ , and  $\text{tr}(M(\mathbf{x})) = \text{tr}(\sigma^2(\hat{\mathbf{x}})) = a^2 + (d-1)b^2 > 0$  for all  $\mathbf{x} \in \mathbb{R}^d$ .

*Remark 3.1.* The seeming ambiguity in the definition of  $\Delta$  due to the choice of  $Q_{\hat{\mathbf{x}}}$  can be resolved by noting that  $\Delta$  can be rewritten as

$$\Delta = Q_{\hat{\mathbf{x}}} D Q_{\hat{\mathbf{x}}}^\top Q_{\hat{\mathbf{x}}} \zeta = Q_{\hat{\mathbf{x}}} D Q_{\hat{\mathbf{x}}}^\top \tilde{\zeta},$$

where  $\tilde{\zeta} = Q_{\hat{\mathbf{x}}} \zeta$  is also uniform on  $\mathbb{S}^{d-1}$  (this follows from the spherical symmetry of the uniform distribution on  $\mathbb{S}^{d-1}$ ). Moreover, the symmetric matrix  $H_{\hat{\mathbf{x}}} := Q_{\hat{\mathbf{x}}} D Q_{\hat{\mathbf{x}}}^\top$  is determined explicitly in terms of  $\hat{\mathbf{x}}$ :

$$\begin{aligned} H_{\hat{\mathbf{x}}} &= Q_{\hat{\mathbf{x}}} D Q_{\hat{\mathbf{x}}}^\top = Q_{\hat{\mathbf{x}}} (b\sqrt{d}I + (a-b)\sqrt{d}\mathbf{e}_1 \mathbf{e}_1^\top) Q_{\hat{\mathbf{x}}}^\top \\ &= b\sqrt{d}I + (a-b)\sqrt{d}\hat{\mathbf{x}}\hat{\mathbf{x}}^\top. \end{aligned}$$

Consequently, we could choose to specify  $\Delta$  explicitly as

$$\Delta = H_{\hat{\mathbf{x}}} \tilde{\zeta} = b\sqrt{d}\tilde{\zeta} + (a-b)\sqrt{d}\hat{\mathbf{x}}\langle \hat{\mathbf{x}}, \tilde{\zeta} \rangle,$$

with  $\tilde{\zeta}$  taken to be uniform on  $\mathbb{S}^{d-1}$ . As before, we find that  $\mathbb{E}_{\mathbf{x}}[\Delta] = H_{\hat{\mathbf{x}}} \mathbb{E}[\tilde{\zeta}] = \mathbf{0}$  and

$$\mathbb{E}_{\mathbf{x}}[\Delta \Delta^\top] = H_{\hat{\mathbf{x}}} \mathbb{E}[\tilde{\zeta} \tilde{\zeta}^\top] H_{\hat{\mathbf{x}}} = \frac{1}{d} H_{\hat{\mathbf{x}}}^2 = a^2 \hat{\mathbf{x}} \hat{\mathbf{x}}^\top + b^2 (I - \hat{\mathbf{x}} \hat{\mathbf{x}}^\top).$$

Recall that we assume our random walk to be time-homogeneous, so that equation (3.1) in fact determines the distribution of  $\Delta_n$  for all  $n \geq 0$ . Formally, we define  $\zeta_0, \zeta_1, \dots$  a sequence of independent random variables uniformly distributed on  $\mathbb{S}^{d-1}$ , and for each  $n \geq 0$  we define  $\Delta_n$  conditional on  $\{X_n = \mathbf{x}\}$  via

$$\Delta_n = Q_{\hat{\mathbf{x}}} D \zeta_n. \tag{3.2}$$

We call  $X = (X_n, n \in \mathbb{Z}_+)$  defined in this way an *elliptic random walk*.

As a corollary to Theorems 2.3 and 2.4, we get the following recurrence classification for the elliptic random walk model. For this model the  $\varepsilon(r)$  in (A3) is identically zero so we get a complete classification that includes the boundary case.

**Corollary 3.2.** *Let  $d \geq 2$  and  $a, b \in (0, \infty)$ . Let  $X$  be an elliptic random walk on  $\mathbb{R}^d$ . Then  $X$  is transient if  $a^2 < (d-1)b^2$  and null-recurrent if  $a^2 \geq (d-1)b^2$ .*

In two dimensions we can explicitly describe the random walk as follows. For  $\mathbf{x} \in \mathbb{R}^2$ ,  $\mathbf{x} \neq \mathbf{0}$  with  $\mathbf{x} = (x_1, x_2)$  in Cartesian components, set  $\mathbf{x}^\perp := (-x_2, x_1)$ . Fix  $a, b \in (0, \infty)$ . Let  $E_{\mathbf{x}}(a, b)$  denote the ellipse with centre  $\mathbf{x}$  and principal axes aligned in the  $\mathbf{x}$ ,  $\mathbf{x}^\perp$  directions, with lengths  $2\sqrt{2}a$ ,  $2\sqrt{2}b$  respectively, given in parametrized form by

$$E_{\mathbf{x}}(a, b) := \left\{ \mathbf{x} + \sqrt{2}a \frac{\mathbf{x}}{\|\mathbf{x}\|} \cos \phi + \sqrt{2}b \frac{\mathbf{x}^\perp}{\|\mathbf{x}\|} \sin \phi : \phi \in (-\pi, \pi] \right\}, \quad (3.3)$$

and for  $\mathbf{x} = \mathbf{0}$  set

$$E_{\mathbf{0}}(a, b) := \left\{ \sqrt{2}a \mathbf{e}_1 \cos \phi + \sqrt{2}b \mathbf{e}_2 \sin \phi : \phi \in (-\pi, \pi] \right\}.$$

The parameter  $\phi$  in the parametrization (3.3) should be interpreted with caution: it is *not*, in general, the central angle of the parametrized point on the ellipse.

Given  $X_n = \mathbf{x} \in \mathbb{R}^2$ ,  $X_{n+1}$  is taken to be distributed on  $E_{\mathbf{x}}(a, b)$ , ‘uniformly’ with respect to the parametrization (3.3). Precisely, let  $\phi_0, \phi_1, \dots$  be a sequence of independent random variables uniformly distributed on  $(-\pi, \pi]$ . Then, on  $\{X_n \neq \mathbf{0}\}$ ,

$$X_{n+1} = X_n + \sqrt{2}a \frac{X_n}{\|X_n\|} \cos \phi_n + \sqrt{2}b \frac{X_n^\perp}{\|X_n\|} \sin \phi_n, \quad (3.4)$$

while, on  $\{X_n = \mathbf{0}\}$ ,

$$X_{n+1} = (\sqrt{2}a \cos \phi_n, \sqrt{2}b \sin \phi_n). \quad (3.5)$$

Figure 4 shows two sample paths of a simulation of the elliptic random walk in  $\mathbb{R}^2$  in the two cases of recurrence and transience. In each picture the walk starts at the origin at the centre of the picture; time is represented by the variation in colour (from red to yellow, or from dark to light if viewed in grey-scale).

*Remarks 3.3.* (a) The process  $X$  reduces to the classical PRRW when  $a = b$ : in that case it is spatially homogeneous, i.e., the distribution of the increment  $X_{n+1} - X_n$  does not depend on  $X_n$ . For  $a \neq b$  the random walk is not spatially homogeneous, and the jump distribution depends upon the projection onto the unit sphere of the walk’s current position.

(b) As mentioned earlier, we choose to take increments as defined at (3.4), rather than increments that are uniform on the ellipse with respect to one-dimensional Lebesgue measure on  $E_{\mathbf{x}}(a, b)$ , purely for computational reasons. In fact, in two dimensions, since the Lebesgue measure on  $E_{\mathbf{x}}(a, b)$  coincides with the measure induced by taking  $\phi$  uniformly distributed on  $(-\pi, \pi]$  when  $a = b$ , and the case  $a = b$  is critically recurrent, the qualitative behaviour will be the same in either case: the walk will be transient for  $a < b$  and recurrent for  $a \geq b$ . For higher dimensions, taking increments that are uniform with respect to the Lebesgue measure on  $E_{\mathbf{x}}^d(a, b) := \{Q_{\mathbf{x}} D \mathbf{u} : \mathbf{u} \in \mathbb{S}^{d-1}\}$  will still specify a family of models that exhibit a phase transition, from transience (for  $a/b$  small) to recurrence (for  $a/b$  large) but the exact shape of the ellipsoid in the critical case (i.e., the smallest ratio  $a/b$  for which the walk is recurrent) may be different.



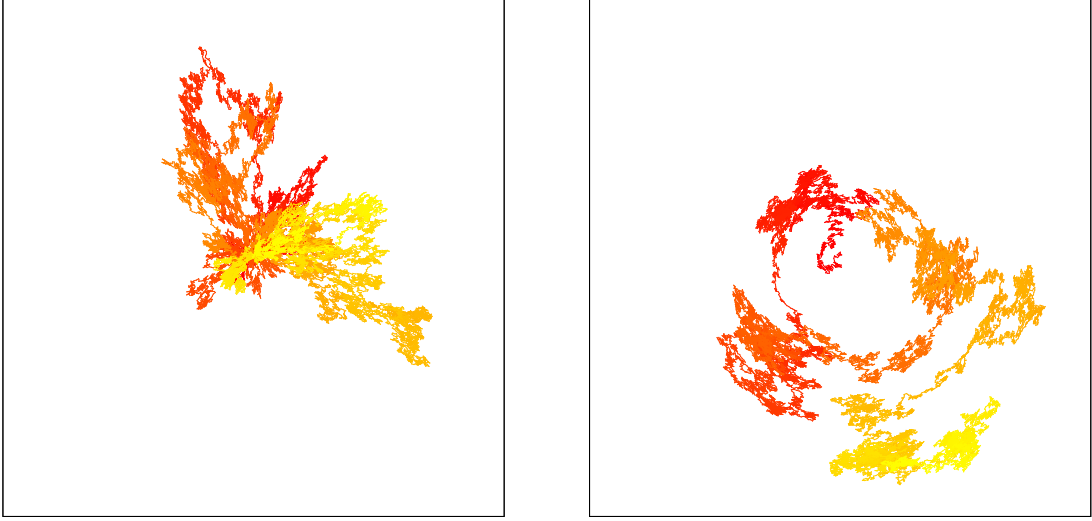


Figure 4: Simulation of the elliptic random walk in  $\mathbb{R}^2$  for the recurrent case  $a > b$  (*left*) and the transient case  $a < b$  (*right*).

(c) It follows from (3.2) that

$$\begin{aligned}
\|X_{n+1}\|^2 &= \|X_n\|^2 + 2\|X_n\|\langle \widehat{X}_n, \Delta_n \rangle + \|\Delta_n\|^2 \\
&= \|X_n\|^2 + 2\|X_n\|\langle \mathbf{e}_1, D\zeta_n \rangle + \langle \zeta_n, D^2\zeta_n \rangle \\
&= \|X_n\|^2 + 2a\sqrt{d}\|X_n\|\langle \mathbf{e}_1, \zeta_n \rangle + (a^2 - b^2)d\langle \mathbf{e}_1, \zeta_n \rangle^2 + b^2d.
\end{aligned} \tag{3.6}$$

In particular, for this family of models ( $\|X_n\|, n \in \mathbb{Z}_+$ ) is itself a Markov process, since the distribution of  $\|X_{n+1}\|$  depends only on  $\|X_n\|$  and not  $X_n$ ; however, in the general setting of Section 2, this need not be the case.

One-dimensional processes with evolutions reminiscent to that given by (3.6) have been studied previously by Kingman [11] and Bingham [2]. Those processes can be viewed, respectively, as the distance from its start point of a random walk in Euclidean space, and the geodesic distance from its start point of a random walk on the surface of a sphere, but in both cases the increments of the random walk have the property that the distribution of the jump vector is a product of the independent marginal distributions of the length and direction of the jump vector. In contrast, for the elliptic random walk the laws of  $\|\Delta_n\|$  and  $\langle \widehat{X}_n, \widehat{\Delta}_n \rangle$  are *not* independent (except when  $a = b$ ).

(d) The theory equally applies to the case where the ellipsoid specifying the jump distribution is oriented with some fixed angle  $\alpha \in [0, \pi)$  with respect to the radial direction. If we define  $\Delta = Q_{\hat{\mathbf{x}}}^\alpha D\zeta$ , where  $Q_{\hat{\mathbf{x}}}^\alpha$  is an orthogonal matrix that maps  $\mathbf{e}_\alpha := \mathbf{e}_1 \cos \alpha + \mathbf{e}_2 \sin \alpha$  to  $\hat{\mathbf{x}}$ , then we find that transience of  $X$  is equivalent to

$$(a^2 - b^2) \cos 2\alpha < (d - 2)b^2.$$

Note that for  $d = 2$ ,  $Q_{\hat{\mathbf{x}}}^\alpha$  and therefore  $\Delta$  are well defined, but this is not so for higher dimensions. Nevertheless, for *any* collection of matrices  $(Q_{\mathbf{u}}^\alpha; \mathbf{u} \in \mathbb{S}^{d-1})$  satisfying

$Q_{\mathbf{u}}^\alpha \mathbf{e}_\alpha = \mathbf{u}$  for all  $\mathbf{u} \in \mathbb{S}^{d-1}$  we get the same recurrence classification. This is because the distribution of  $\|X_{n+1}\|$  given  $X_n$  is determined through the angle  $\alpha$  via

$$\|X_n\|^2 + 2\sqrt{d}\|X_n\|(a\langle \mathbf{e}_1, \boldsymbol{\zeta}_n \rangle \cos \alpha + b\langle \mathbf{e}_2, \boldsymbol{\zeta}_n \rangle \sin \alpha) + (a^2 - b^2)d\langle \mathbf{e}_1, \boldsymbol{\zeta}_n \rangle^2 + b^2d,$$

and therefore assumption (A4) holds with  $U = a^2 \cos^2 \alpha + b^2 \sin^2 \alpha$  and  $V = a^2 + (d-1)b^2$ .

## 4 Non-confinement

In this section we prove that the assumptions (A0), (A1), and (A2) imply that  $\limsup_{n \rightarrow \infty} \|X_n\| = +\infty$ , a.s. We first present a general result for martingales on  $\mathbb{R}^d$  satisfying a “uniform dispersion” condition; the result can be viewed as a  $d$ -dimensional martingale version of *Kolmogorov’s other inequality* (see e.g. [5, pp. 123, 502]).

**Lemma 4.1.** *Let  $d \in \mathbb{N}$ . Suppose that  $(Y_n, n \in \mathbb{Z}_+)$  is an  $\mathbb{R}^d$ -valued process adapted to a filtration  $(\mathcal{G}_n, n \in \mathbb{Z}_+)$ , with  $\mathbb{P}[Y_0 = \mathbf{0} \mid \mathcal{G}_0] = 1$ . Suppose that there exist  $p > 2, v > 0, B < \infty$  such that for all  $n \in \mathbb{Z}_+$ , a.s.,*

$$\mathbb{E}[\|Y_{n+1} - Y_n\|^p \mid \mathcal{G}_n] \leq B; \quad (4.1)$$

$$\mathbb{E}[\|Y_{n+1} - Y_n\|^2 \mid \mathcal{G}_n] \geq v; \quad (4.2)$$

$$\mathbb{E}[Y_{n+1} - Y_n \mid \mathcal{G}_n] = \mathbf{0}. \quad (4.3)$$

Then there exists  $D < \infty$ , depending only on  $B, p$ , and  $v$ , such that for all  $n \in \mathbb{Z}_+$  and all  $x \in \mathbb{R}_+$ ,

$$\mathbb{P}\left[\max_{0 \leq \ell \leq n} \|Y_\ell\| \geq x \mid \mathcal{G}_0\right] \geq 1 - \frac{D(1+x)^2}{n}, \text{ a.s.}$$

*Proof.* Let  $x > 0$  and set  $\tau = \min\{n \geq 0 : \|Y_n\| \geq x\}$ ; throughout the paper we adopt the usual convention  $\min \emptyset := \infty$ . In analogy with previous notation, write  $\Delta_n = Y_{n+1} - Y_n$  for the jump distribution, and let

$$W_n = \begin{cases} Y_n & \text{if } \|Y_n\| \leq A(1+x), \\ Y_{n-1} + \widehat{\Delta}_{n-1}(A-1)(1+x) & \text{if } \|Y_n\| > A(1+x), \end{cases}$$

where  $A > 1$  is a constant to be specified later. Note that  $W_n$  is  $\mathcal{G}_n$ -measurable.

Now, on  $\{\|Y_n\| \leq x\}$ ,  $W_n = Y_n$  and

$$\begin{aligned} \mathbb{E}[W_{n+1} - W_n \mid \mathcal{G}_n] &= \mathbb{E}[\Delta_n \mid \mathcal{G}_n] \\ &\quad + \mathbb{E}[\widehat{\Delta}_n ((A-1)(1+x) - \|\Delta_n\|) \mathbf{1}\{\|Y_{n+1}\| > A(1+x)\} \mid \mathcal{G}_n]. \end{aligned}$$

But  $\{\|Y_{n+1}\| > A(1+x)\} \cap \{\|Y_n\| \leq x\}$  implies that  $\|\Delta_n\| > (A-1)(1+x)$ , and by (4.3),  $\mathbb{E}[\Delta_n \mid \mathcal{G}_n] = \mathbf{0}$ . Hence, on  $\{\|Y_n\| \leq x\}$ ,

$$\begin{aligned} \|\mathbb{E}[W_{n+1} - W_n \mid \mathcal{G}_n]\| &\leq \mathbb{E}[\|\Delta_n\| \mathbf{1}\{\|\Delta_n\| > (A-1)(1+x)\} \mid \mathcal{G}_n] \\ &\leq (A-1)^{-1}(1+x)^{-1} \mathbb{E}[\|\Delta_n\|^2 \mid \mathcal{G}_n] \\ &\leq B'(A-1)^{-1}(1+x)^{-1}, \text{ a.s.,} \end{aligned}$$

where, by (4.1) and Lyapunov's inequality,  $B' < \infty$  depends only on  $B$  and  $p$ . Hence we can choose  $A \geq A_0$  for some  $A_0 = A_0(B, p, v)$  large enough so that

$$\|\mathbb{E}[W_{n+1} - W_n \mid \mathcal{G}_n]\| \leq (v/8)(1+x)^{-1}, \text{ on } \{\|Y_n\| \leq x\}. \quad (4.4)$$

Also, on  $\{\|Y_n\| \leq x\}$ , by a similar argument,

$$\begin{aligned} \mathbb{E}[\|W_{n+1} - W_n\|^2 \mid \mathcal{G}_n] &= \mathbb{E}[\|\Delta_n\|^2 \mid \mathcal{G}_n] \\ &\quad + \mathbb{E}[(A-1)^2(1+x)^2 - \|\Delta_n\|^2] \mathbf{1}\{\|Y_{n+1}\| > A(1+x)\} \mid \mathcal{G}_n] \\ &\geq \mathbb{E}[\|\Delta_n\|^2 \mid \mathcal{G}_n] - \mathbb{E}[\|\Delta_n\|^2 \mathbf{1}\{\|\Delta_n\| > (A-1)(1+x)\} \mid \mathcal{G}_n] \\ &\geq v - (A-1)^{2-p}(1+x)^{2-p} \mathbb{E}[\|\Delta_n\|^p \mid \mathcal{G}_n] \\ &\geq v/2, \end{aligned} \quad (4.5)$$

for all  $x \geq 0$  and  $A \geq A_1$  for sufficiently large  $A_1 = A_1(B, p, v)$ , using (4.1) and (4.2).

Now, set  $Z_n = \|W_{n \wedge \tau}\|^2$ . Then, on  $\{n < \tau\}$ , by (4.4) and (4),

$$\begin{aligned} \mathbb{E}[Z_{n+1} - Z_n \mid \mathcal{G}_n] &= \mathbb{E}[\|W_{n+1}\|^2 - \|W_n\|^2 \mid \mathcal{G}_n] \\ &= \mathbb{E}[\|W_{n+1} - W_n\|^2 \mid \mathcal{G}_n] + 2\langle W_n, \mathbb{E}[W_{n+1} - W_n \mid \mathcal{G}_n] \rangle \\ &\geq \frac{v}{2} - \frac{2\|W_n\|v}{8(1+x)} \geq \frac{v}{2} - \frac{vx}{4(1+x)} \geq \frac{v}{4}. \end{aligned}$$

Hence  $Z_n - \sum_{k=0}^{n-1} v_k$  is a  $\mathcal{G}_n$ -adapted submartingale, where

$$v_k = \frac{v}{4} \mathbf{1}\{k < \tau\} \geq \frac{v}{4} \mathbf{1}\{n < \tau\}, \text{ for } 0 \leq k < n.$$

By construction,  $0 \leq Z_n \leq A^2(1+x)^2$ , so

$$0 = \mathbb{E}[Z_0 \mid \mathcal{G}_0] \leq \mathbb{E}[Z_n \mid \mathcal{G}_0] - \sum_{k=0}^{n-1} \mathbb{E}[v_k \mid \mathcal{G}_0] \leq A^2(1+x)^2 - \sum_{k=0}^{n-1} \frac{v}{4} \mathbb{P}[n < \tau \mid \mathcal{G}_0],$$

which implies  $n(v/4)\mathbb{P}[n < \tau \mid \mathcal{G}_0] \leq A^2(1+x)^2$ . In other words,

$$\mathbb{P}\left[\max_{0 \leq \ell \leq n} \|Y_\ell\| < x \mid \mathcal{G}_0\right] \leq \frac{4A^2(1+x)^2}{vn}, \text{ a.s.} \quad \square$$

Now we can give the proof of Proposition 2.1.

*Proof of Proposition 2.1.* It is enough to show that for all  $x \in \mathbb{R}_+$  the event  $\{\|X_n\| \geq x\}$  occurs infinitely often. For a given  $x$ , we will apply Lemma 4.1 to  $Y_n = X_{m+n} - X_m$  with  $\mathcal{G}_n = \sigma(X_0, \dots, X_{m+n})$ ; that result is applicable, since (A0), (A1) and (A2) imply (4.1), (4.3) and (4.2), respectively. Thus Lemma 4.1 shows that, for some finite  $t = t(x)$ ,

$$\mathbb{P}\left[\max_{0 \leq \ell \leq t-1} \|X_{m+\ell} - X_m\| \geq 2x \mid X_0, \dots, X_m\right] \geq \frac{1}{2}, \text{ a.s.}, \quad (4.6)$$

for all  $m \geq 0$ . For  $k = 1, 2, \dots$ , define the event

$$A_k = \left\{ \max_{0 \leq \ell \leq t-1} \|X_{(k-1)t+\ell} - X_{(k-1)t}\| \geq 2x \right\},$$

and filtration  $\mathcal{G}'_{k-1} = \sigma(X_0, \dots, X_{(k-1)t})$ . Then  $A_k \in \mathcal{G}'_k$ , and, by (4.6),  $\mathbb{P}[A_k \mid \mathcal{G}'_{k-1}] \geq \frac{1}{2}$ , a.s., for all  $k$ . An application of Lévy's extension of the Borel–Cantelli lemma (see, e.g., [9, Cor. 7.20]) shows that  $A_k$  occurs infinitely often, a.s. For each  $k$  such that  $A_k$  occurs, either

- $\|X_{(k-1)t}\| \geq x$ , or
- $\|X_{(k-1)t}\| \leq x$  and  $\|X_n\| \geq x$  for some  $(k-1)t < n < kt$ .

Since one of these cases must occur for infinitely many  $k$ , we have that  $\{\|X_n\| \geq x\}$  occurs infinitely often, as required.  $\square$

## 5 Recurrence classification

In this section we study the random walk  $X_n$  and give the proof of the recurrence classification, Theorem 2.3. The method of proof is based on applying classical results of Lamperti [12] to the  $\mathbb{R}_+$ -valued radial process given by  $R_n := \|X_n\|$ . The method rests on an analysis of the increments  $R_{n+1} - R_n$  given  $X_n = \mathbf{x} \in \mathbb{X}$ ; in general,  $R_n$  is not itself a Markov process. The following notation will be useful. Given  $\mathbf{x} \neq \mathbf{0}$  and  $\mathbf{y} \in \mathbb{R}^d$ , write

$$\mathbf{y}_{\mathbf{x}} := \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|} = \langle \hat{\mathbf{x}}, \mathbf{y} \rangle,$$

so that  $\mathbf{y}_{\mathbf{x}}$  is the component of  $\mathbf{y}$  in the  $\hat{\mathbf{x}}$  direction, and  $\mathbf{y} - \mathbf{y}_{\mathbf{x}} \hat{\mathbf{x}}$  is a vector perpendicular to  $\hat{\mathbf{x}}$ .

First we state a general result on the increments of  $R_n$  for a Markov process  $X_n$  on  $\mathbb{X}$ . Recall that we write  $\Delta = X_1 - X_0$ , and let  $\Delta_{\mathbf{x}}$  be the radial component of  $\Delta$  at  $X_0 = \mathbf{x}$  in accordance with the notation described above; no confusion should arise with our notation  $\Delta_n$  defined previously.

We make an important comment on notation. When we write  $O(\|\mathbf{x}\|^{-1-\delta})$ , and similar expressions, these are understood to be uniform in  $\mathbf{x}$ . That is, if  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  and  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , we write  $f(\mathbf{x}) = O(g(\|\mathbf{x}\|))$  to mean that there exist  $C \in \mathbb{R}_+$  and  $r \in \mathbb{R}_+$  such that

$$|f(\mathbf{x})| \leq Cg(\|\mathbf{x}\|) \text{ for all } \mathbf{x} \in \mathbb{X} \text{ with } \|\mathbf{x}\| \geq r. \quad (5.1)$$

**Lemma 5.1.** *Suppose that  $X$  is a discrete-time, time-homogeneous Markov process on  $\mathbb{X} \subseteq \mathbb{R}^d$  satisfying (A0) for some  $p > 2$ . Then, for  $R_n := \|X_n\|$ , we have*

$$\sup_{\mathbf{x} \in \mathbb{X}} \mathbb{E}[|R_{n+1} - R_n|^p \mid X_n = \mathbf{x}] < \infty, \quad (5.2)$$

and the radial increment moment functions satisfy

$$\mu_1(\mathbf{x}) := \mathbb{E}[R_{n+1} - R_n \mid X_n = \mathbf{x}] = \mathbb{E}_{\mathbf{x}}[\Delta_{\mathbf{x}}] + \frac{\mathbb{E}_{\mathbf{x}}[\|\Delta\|^2 - \Delta_{\mathbf{x}}^2]}{2\|\mathbf{x}\|} + O(\|\mathbf{x}\|^{-1-\delta}), \quad (5.3)$$

$$\mu_2(\mathbf{x}) := \mathbb{E}[(R_{n+1} - R_n)^2 \mid X_n = \mathbf{x}] = \mathbb{E}_{\mathbf{x}}[\Delta_{\mathbf{x}}^2] + O(\|\mathbf{x}\|^{-\delta}), \quad (5.4)$$

as  $\|\mathbf{x}\| \rightarrow \infty$ , for some  $\delta = \delta(p) > 0$ .

*Proof.* By time-homogeneity, it suffices to consider the case  $n = 0$ . By the triangle inequality,  $|R_1 - R_0| = |\|X_0 + \Delta\| - \|X_0\|| \leq \|\Delta\|$ , so that (5.2) follows from (A0).

We prove (5.3) and (5.4) by approximating

$$\begin{aligned} \|\mathbf{x} + \Delta\| - \|\mathbf{x}\| &= \sqrt{\langle \mathbf{x} + \Delta, \mathbf{x} + \Delta \rangle} - \|\mathbf{x}\| \\ &= \|\mathbf{x}\| \left[ \left( 1 + \frac{2\Delta_{\mathbf{x}}}{\|\mathbf{x}\|} + \frac{\|\Delta\|^2}{\|\mathbf{x}\|^2} \right)^{1/2} - 1 \right] \end{aligned} \quad (5.5)$$

for large  $\mathbf{x}$ . Let  $A_{\mathbf{x}} = \{\|\Delta\| \leq \|\mathbf{x}\|^\beta\}$  for some  $\beta \in (0, 1)$  to be determined later. On the event  $A_{\mathbf{x}}$  we approximate (5.5) using Taylor's formula for  $(1+y)^{1/2}$ , and on the event  $A_{\mathbf{x}}^c$  we bound (5.5) using (A0).

Indeed, for all  $y > -1$ , Taylor's theorem with Lagrange remainder shows that

$$(1+y)^{1/2} = 1 + \frac{1}{2}y - \frac{1}{8}y^2(1+\gamma y)^{-3/2},$$

for some  $\gamma = \gamma(y) \in [0, 1]$ , so on the event  $A_{\mathbf{x}}$ ,

$$\begin{aligned} \|\mathbf{x} + \Delta\| - \|\mathbf{x}\| &= \|\mathbf{x}\| \left( \frac{\Delta_{\mathbf{x}}}{\|\mathbf{x}\|} + \frac{\|\Delta\|^2}{2\|\mathbf{x}\|^2} - \frac{1}{8} \left( \frac{2\Delta_{\mathbf{x}}}{\|\mathbf{x}\|} + \frac{\|\Delta\|^2}{\|\mathbf{x}\|^2} \right)^2 (1 + O(\|\mathbf{x}\|^{\beta-1})) \right) \\ &= \Delta_{\mathbf{x}} + \frac{\|\Delta\|^2}{2\|\mathbf{x}\|} - \frac{\|\mathbf{x}\|}{8} \left( \frac{4\Delta_{\mathbf{x}}^2}{\|\mathbf{x}\|^2} + \frac{\|\Delta\|^2}{\|\mathbf{x}\|^2} \left( \frac{4\Delta_{\mathbf{x}}}{\|\mathbf{x}\|} + \frac{\|\Delta\|^2}{\|\mathbf{x}\|^2} \right) \right) (1 + O(\|\mathbf{x}\|^{\beta-1})) \\ &= \Delta_{\mathbf{x}} + \frac{\|\Delta\|^2}{2\|\mathbf{x}\|} - \frac{\Delta_{\mathbf{x}}^2}{2\|\mathbf{x}\|} (1 + O(\|\mathbf{x}\|^{\beta-1})) + \frac{\|\Delta\|^2}{\|\mathbf{x}\|} O(\|\mathbf{x}\|^{\beta-1}) \\ &= \Delta_{\mathbf{x}} + \left( \frac{\|\Delta\|^2 - \Delta_{\mathbf{x}}^2}{2\|\mathbf{x}\|} \right) (1 + O(\|\mathbf{x}\|^{\beta-1})), \end{aligned} \quad (5.6)$$

where the error terms follow from the fact that  $|\Delta_{\mathbf{x}}| \leq \|\Delta\| \leq \|\mathbf{x}\|^\beta$  for  $\beta < 1$ .

On the other hand,

$$\|\|\mathbf{x} + \Delta\| - \|\mathbf{x}\|\| \mathbf{1}(A_{\mathbf{x}}^c) \leq \|\Delta\| \mathbf{1}(A_{\mathbf{x}}^c) = \|\Delta\|^p \|\Delta\|^{1-p} \mathbf{1}(A_{\mathbf{x}}^c) \leq \|\Delta\|^p \|\mathbf{x}\|^{\beta(1-p)}, \quad (5.7)$$

by the triangle inequality and the fact that  $\|\Delta\| > \|\mathbf{x}\|^\beta$  on  $A_{\mathbf{x}}^c$ . Since

$$\|\mathbf{x} + \Delta\| - \|\mathbf{x}\| = (\|\mathbf{x} + \Delta\| - \|\mathbf{x}\|) \mathbf{1}(A_{\mathbf{x}}) + (\|\mathbf{x} + \Delta\| - \|\mathbf{x}\|) \mathbf{1}(A_{\mathbf{x}}^c),$$

we can combine (5.6) and (5.7) to give

$$\begin{aligned} &\left| \|\mathbf{x} + \Delta\| - \|\mathbf{x}\| - \left[ \Delta_{\mathbf{x}} + \left( \frac{\|\Delta\|^2 - \Delta_{\mathbf{x}}^2}{2\|\mathbf{x}\|} \right) (1 + O(\|\mathbf{x}\|^{\beta-1})) \right] \right| \\ &= \left| \|\mathbf{x} + \Delta\| - \|\mathbf{x}\| - \left[ \Delta_{\mathbf{x}} + \left( \frac{\|\Delta\|^2 - \Delta_{\mathbf{x}}^2}{2\|\mathbf{x}\|} \right) (1 + O(\|\mathbf{x}\|^{\beta-1})) \right] \right| \mathbf{1}(A_{\mathbf{x}}^c) \\ &\leq \|\Delta\|^p \|\mathbf{x}\|^{\beta(1-p)} + \left| \Delta_{\mathbf{x}} + \left( \frac{\|\Delta\|^2 - \Delta_{\mathbf{x}}^2}{2\|\mathbf{x}\|} \right) (1 + O(\|\mathbf{x}\|^{\beta-1})) \right| \mathbf{1}(A_{\mathbf{x}}^c) \\ &\leq 2\|\Delta\|^p \|\mathbf{x}\|^{\beta(1-p)} + \frac{\|\Delta\|^p}{2\|\mathbf{x}\|} (1 + O(\|\mathbf{x}\|^{\beta-1})) \|\mathbf{x}\|^{\beta(2-p)}. \end{aligned}$$

Therefore, taking expectations and using (A0), we obtain

$$\mu_1(\mathbf{x}) = \mathbb{E}_{\mathbf{x}}[\Delta_{\mathbf{x}}] + \frac{\mathbb{E}_{\mathbf{x}}[\|\Delta\|^2 - \Delta_{\mathbf{x}}^2]}{2\|\mathbf{x}\|} + O(\|\mathbf{x}\|^{\beta-2}) + O(\|\mathbf{x}\|^{\beta(1-p)}) + O(\|\mathbf{x}\|^{\beta(2-p)-1}).$$

Taking  $\beta = 2/p$  makes all the error terms of size  $O(\|\mathbf{x}\|^{-1-\delta})$  for some  $\delta = \delta(p) > 0$ , namely for  $\delta = (p-2)/p$ .

For the second moment, we use the identity

$$\begin{aligned} (\|\mathbf{x} + \Delta\| - \|\mathbf{x}\|)^2 &= \|\mathbf{x} + \Delta\|^2 - \|\mathbf{x}\|^2 - 2\|\mathbf{x}\|(\|\mathbf{x} + \Delta\| - \|\mathbf{x}\|) \\ &= 2\|\mathbf{x}\|\Delta_{\mathbf{x}} + \|\Delta\|^2 - 2\|\mathbf{x}\|(\|\mathbf{x} + \Delta\| - \|\mathbf{x}\|), \end{aligned}$$

so that

$$\mu_2(\mathbf{x}) = 2\|\mathbf{x}\| \mathbb{E}_{\mathbf{x}}[\Delta_{\mathbf{x}}] + \mathbb{E}_{\mathbf{x}}[\|\Delta\|^2] - 2\|\mathbf{x}\|\mu_1(\mathbf{x}) = \mathbb{E}_{\mathbf{x}}[\Delta_{\mathbf{x}}^2] + O(\|\mathbf{x}\|^{-\delta}),$$

as required.  $\square$

With the additional assumptions (A1), (A3), and (A4), we can use Lemma 5.1 to prove the following result.

**Lemma 5.2.** *Suppose that  $X$  is a discrete-time, time-homogeneous Markov process on  $\mathbb{X} \subseteq \mathbb{R}^d$  satisfying (A0), (A1), (A3), and (A4). Then, with  $\mu_1, \mu_2$  defined at (5.3), (5.4), and  $\varepsilon(r)$  defined at (A3), there exists  $\delta > 0$  such that, as  $\|\mathbf{x}\| \rightarrow \infty$ ,*

$$2\|\mathbf{x}\|\mu_1(\mathbf{x}) = V - U + O(\varepsilon(\|\mathbf{x}\|)) + O(\|\mathbf{x}\|^{-\delta}), \quad \mu_2(\mathbf{x}) = U + O(\varepsilon(\|\mathbf{x}\|)) + O(\|\mathbf{x}\|^{-\delta}). \quad (5.8)$$

*Proof.* By definition of  $\varepsilon(r)$  at (A3) we have  $\|M(\mathbf{x}) - \sigma^2(\hat{\mathbf{x}})\|_{\text{op}} = O(\varepsilon(\|\mathbf{x}\|))$  as  $\|\mathbf{x}\| \rightarrow \infty$ . Then (A4) implies that

$$\begin{aligned} \mathbb{E}_{\mathbf{x}}[\|\Delta\|^2] &= \text{tr}(M(\mathbf{x})) \\ &= \text{tr}(\sigma^2(\hat{\mathbf{x}})) + O(\varepsilon(\|\mathbf{x}\|)) \\ &= V + O(\varepsilon(\|\mathbf{x}\|)), \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}_{\mathbf{x}}[\Delta_{\mathbf{x}}^2] &= \langle \hat{\mathbf{x}}, M(\mathbf{x}) \cdot \hat{\mathbf{x}} \rangle \\ &= \langle \hat{\mathbf{x}}, \sigma^2(\hat{\mathbf{x}}) \cdot \hat{\mathbf{x}} \rangle + O(\varepsilon(\|\mathbf{x}\|)) \\ &= U + O(\varepsilon(\|\mathbf{x}\|)), \end{aligned}$$

and (A1) implies that  $\mathbb{E}_{\mathbf{x}}[\Delta_{\mathbf{x}}] = \mathbb{E}_{\mathbf{x}}[\langle \Delta, \hat{\mathbf{x}} \rangle] = \langle \mu(\mathbf{x}), \hat{\mathbf{x}} \rangle = 0$ . Using these expressions in Lemma 5.1 yields (5.8).  $\square$

Now we can complete the proof of Theorem 2.3.

*Proof of Theorem 2.3.* We apply Lamperti's [12] recurrence classification to  $R_n = \|X_n\|$ , the radial process. Proposition 2.1 shows that  $\limsup_{n \rightarrow \infty} R_n = +\infty$ , and Lemma 5.1 tells us that (5.2) is satisfied.

Because the error terms in (5.8) are uniform in  $\mathbf{x}$ , Lemma 5.2 shows that for all  $\eta > 0$  there exists  $C < \infty$  such that

$$2\|\mathbf{x}\|\mu_1(\mathbf{x}) - \mu_2(\mathbf{x}) \in [V - 2U - \eta, V - 2U + \eta]$$

for all  $\mathbf{x} \in \mathbb{X}$  with  $\|\mathbf{x}\| \geq C$ . Therefore, it follows from Theorem 3.2 of [12] that  $X$  is transient if  $V - 2U > 0$  and recurrent if  $V - 2U < 0$ . For the boundary case, when  $V - 2U = 0$ , if  $\varepsilon(r) = O(r^{-\delta_0})$  then

$$2\|\mathbf{x}\|\mu_1(\mathbf{x}) - \mu_2(\mathbf{x}) = O(\|\mathbf{x}\|^{-\delta_1}),$$

for  $\delta_1 = \min\{\delta, \delta_0\}$ , which implies that  $X$  is recurrent, again by Theorem 3.2 of [12].  $\square$

## 6 Nullity

In this section we give the proof of Theorem 2.4. In the transient case, this is straightforward.

**Lemma 6.1.** *In case (i) of Theorem 2.3, for any bounded  $A \subset \mathbb{R}^d$ , as  $n \rightarrow \infty$ , the null property (2.2) holds.*

*Proof.* It is sufficient to prove (2.2) in the case where  $A = B_r := \{\mathbf{x} \in \mathbb{X} : \|\mathbf{x}\| \leq r\}$ . In case (i),  $X$  is transient, meaning that  $\|X_n\| \rightarrow \infty$  a.s., so that  $\mathbf{1}\{X_n \in B_r\} \rightarrow 0$ , a.s., for any  $r \in \mathbb{R}_+$ . Hence the Cesàro limit in (2.2) is also 0, a.s., and the  $L^q$  convergence follows from the bounded convergence theorem.  $\square$

It remains to consider cases (ii) and (iii), when  $X$  is recurrent. Thus there exists  $r_0 \in \mathbb{R}_+$  such that  $\liminf_{n \rightarrow \infty} \|X_n\| \leq r_0$ , a.s. Let  $\tau_r := \min\{n \in \mathbb{Z}_+ : X_n \in B_r\}$ . It suffices to take  $A = B_r$ ,  $r > r_0$ , so  $X_n \in B_r$  infinitely often. We make the following claim, whose proof is deferred until the end of this section, which says that if the walk has not yet entered a ball of radius  $R$  (for any  $R > r$  big enough), the time until it reaches the ball of radius  $r$  has tail bounded below as displayed.

**Lemma 6.2.** *In cases (ii) and (iii) of Theorem 2.3, there exists a finite  $r_1 \geq r_0$  such that for any  $r > r_1$  and  $R > r$  there exists a finite positive  $c$  such that*

$$\mathbb{P}[\tau_r \geq n + m \mid X_0, \dots, X_n] \geq cm^{-1/2}, \text{ on } \{n < \tau_R\}, \quad (6.1)$$

for all sufficiently large  $m$ .

Assuming this result, we can complete the proof of Theorem 2.4.

*Proof of Theorem 2.4.* In case (i), the result is contained in Lemma 6.1. So consider cases (ii) and (iii). Fix  $r$  and  $R$  with  $R > r > r_1$ , with  $r_1$  as in Lemma 6.2. Note that  $\liminf_{n \rightarrow \infty} \|X_n\| \leq r_0 \leq r_1$ , a.s.

Set  $\gamma_1 := 0$  and then define recursively, for  $\ell \in \mathbb{N}$ , the stopping times

$$\eta_\ell := \min\{n \geq \gamma_\ell : X_n \notin B_R\}, \quad \gamma_{\ell+1} := \min\{n \geq \eta_\ell : X_n \in B_r\},$$

with the convention that  $\min \emptyset := \infty$ . Since  $r > r_0$  and  $\limsup_{n \rightarrow \infty} \|X_n\| = \infty$  (by Proposition 2.1), for all  $\ell \in \mathbb{N}$  we have  $\eta_\ell < \infty$  and  $\gamma_\ell < \infty$ , a.s., and

$$0 = \gamma_1 < \eta_1 < \gamma_2 < \eta_2 < \dots$$

In particular,  $\lim_{\ell \rightarrow \infty} \gamma_\ell = \lim_{\ell \rightarrow \infty} \eta_\ell = \infty$ , a.s.

We now write  $\mathcal{F}_n := \sigma(X_0, \dots, X_n)$ . We use Lemma 4.1 to show that the process must exit from  $B_R$  rapidly enough. Indeed, if  $\kappa$  is any finite stopping time, set  $Y_n = X_{\kappa+n} - X_\kappa$  and  $\mathcal{G}_n = \mathcal{F}_{\kappa+n}$ . Then the assumptions (A0), (A1) and (A2) show that the hypotheses of Lemma 4.1 are satisfied, since, for example,

$$\mathbb{E}[\|Y_{n+1} - Y_n\|^p \mid \mathcal{G}_n] = \mathbb{E}[\|X_{\kappa+n+1} - X_{\kappa+n}\|^p \mid \mathcal{F}_{\kappa+n}] = \mathbb{E}_{X_{\kappa+n}}[\|\Delta\|^p],$$

by the strong Markov property for  $X$  at the finite stopping time  $\kappa + n$ . In particular, another application of Lemma 4.1, similarly to (4.6), shows that we may choose  $n = n(R) \in \mathbb{N}$  sufficiently large so that

$$\mathbb{P}\left[\max_{0 \leq \ell \leq n(R)} \|X_{\kappa+\ell} - X_\kappa\| \geq 2R \mid \mathcal{F}_\kappa\right] \geq \frac{1}{2}, \text{ a.s.}, \quad (6.2)$$

an event whose occurrence ensures that if  $X_\kappa \in B_R$ , then  $X$  exits  $B_R$  before time  $\kappa + n(R)$ . Fix  $k \in \mathbb{N}$ . Then, an application of (6.2) at stopping time  $\kappa = \gamma_k$  shows that

$$\mathbb{P}[\eta_k - \gamma_k > n(R) \mid \mathcal{F}_{\gamma_k}] \leq \mathbb{P}\left[\max_{0 \leq \ell \leq n(R)} \|X_{\gamma_k + \ell} - X_{\gamma_k}\| < 2R \mid \mathcal{F}_{\gamma_k}\right] \leq \frac{1}{2}, \text{ a.s.}$$

Similarly,

$$\begin{aligned} \mathbb{P}[\eta_k - \gamma_k > 2n(R) \mid \mathcal{F}_{\gamma_k}] &= \mathbb{E}[\mathbf{1}\{\eta_k - \gamma_k > n(R)\} \mathbb{E}[\mathbf{1}\{\eta_k - \gamma_k > 2n(R)\} \mid \mathcal{F}_{\gamma_k + n(R)}] \mid \mathcal{F}_{\gamma_k}] \\ &\leq \frac{1}{2} \mathbb{P}[\eta_k - \gamma_k > n(R) \mid \mathcal{F}_{\gamma_k}] \leq \frac{1}{4}, \end{aligned}$$

this time applying (6.2) at stopping time  $\kappa = \gamma_k + n(R)$  as well. Iterating this argument, it follows that  $\mathbb{P}[\eta_k - \gamma_k > m \cdot n(R) \mid \mathcal{F}_{\gamma_k}] \leq 2^{-m}$ , a.s., for all  $m \in \mathbb{N}$ . From here, it is straightforward to deduce that, for some constant  $C < \infty$ , for any  $k \in \mathbb{N}$ ,

$$\mathbb{E}[\eta_k - \gamma_k \mid \mathcal{F}_{\gamma_k}] \leq C, \text{ a.s.} \quad (6.3)$$

On the other hand, the tail estimate (6.1) implies that

$$\mathbb{P}[\gamma_{k+1} - \eta_k \geq m \mid \mathcal{F}_{\eta_k}] \geq cm^{-1/2}, \text{ a.s.}, \quad (6.4)$$

for  $c > 0$  and all sufficiently large  $m$ .

For any  $n \in \mathbb{N}$ , set  $k(n) := \min\{k \geq 2 : \gamma_k > n\}$ , so that  $\gamma_{k(n)-1} \leq n < \gamma_{k(n)}$  for  $k(n) \in \{2, 3, \dots\}$ . Note  $k(n) < \infty$  and  $\lim_{n \rightarrow \infty} k(n) = \infty$ , a.s. Then we claim

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k \in B_r\} \leq \frac{\sum_{k=1}^{k(n)-1} (\eta_k - \gamma_k)}{\sum_{k=1}^{k(n)-2} (\gamma_{k+1} - \eta_k)}. \quad (6.5)$$

This is easiest to see by considering two separate cases. First, if  $\eta_{k(n)-1} < n < \gamma_{k(n)}$ ,

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k \in B_r\} \leq \frac{1}{\eta_{k(n)-1}} \sum_{k=0}^{\eta_{k(n)-1}-1} \mathbf{1}\{X_k \in B_r\},$$

which implies (6.5), since the set of  $k$  less than  $n$  for which  $X_k \in B_r$  is contained in the set  $\cup_{k=1}^{k(n)-1} [\gamma_k, \eta_k]$ . On the other hand, if  $\gamma_{k(n)-1} \leq n \leq \eta_{k(n)-1}$ , using the elementary inequality  $\frac{a}{b} \leq \frac{a+c}{b+c}$  for non-negative  $a, b, c$  with  $a/b \leq 1$ , we have

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}\{X_k \in B_r\} \leq \frac{1}{\eta_{k(n)-1}} \left( \sum_{k=0}^{n-1} \mathbf{1}\{X_k \in B_r\} + (\eta_{k(n)-1} - n) \right),$$

which again gives (6.5).

To estimate the growth rates of the numerator and denominator of the right-hand side of (6.5), we apply some results from [6]. First, writing  $Z_m = \sum_{k=1}^{m-1} (\eta_k - \gamma_k)$  and  $\mathcal{G}_m = \mathcal{F}_{\gamma_m}$ , by (6.3) we can apply Theorem 2.4 of [6] to the  $\mathcal{G}_m$ -adapted process  $Z_m$  to obtain that for any  $\varepsilon > 0$ , a.s., for all but finitely many  $m$ ,

$$\sum_{k=1}^{m-1} (\eta_k - \gamma_k) \leq m^{1+\varepsilon}.$$



On the other hand, writing  $Z_m = \sum_{k=1}^{m-1} (\gamma_{k+1} - \eta_k)$  and  $\mathcal{G}_m = \mathcal{F}_{\eta_m}$ , by (6.4) we can apply Theorem 2.6 of [6] to the  $\mathcal{G}_m$ -adapted process  $Z_m$  to obtain that for any  $\varepsilon > 0$ , for all  $m$  sufficiently large,

$$\sum_{k=1}^{m-1} (\gamma_{k+1} - \eta_k) \geq m^{2-\varepsilon}.$$

Now (6.5) gives the almost-sure version of the result (2.2). The  $L^q$  version follows from the bounded convergence theorem.  $\square$

It remains to complete the proof of Lemma 6.2. A more general, two-sided version of the inequality in Lemma 6.2 is proved in [7, Theorem 2.4] but under slightly different assumptions. Because of this, we cannot apply that result directly; nevertheless, the proof techniques naturally transfer to our setting. In doing so, the arguments become simpler to apply, so we reproduce them here.

*Proof of Lemma 6.2.* By the Markov property for  $X$  it is enough to prove the statement for  $n = 0$ , namely that there exists finite  $r_1 \geq r_0$  such that for any  $r > r_1$  and  $R > r$  there exists a finite positive constant  $c$  such that, if  $X_0 \notin B_R$  then

$$\mathbb{P}[\tau_r > m \mid X_0] \geq cm^{-1/2},$$

for sufficiently large  $m$ .

We outline the two intuitive steps in the proof. First we show that the probability that  $\max_{0 \leq k \leq \tau_r} \|X_k\|$  exceeds some large  $x$  is bounded below by a constant times  $1/x$ . Second, we show that if the latter event does occur, with probability at least  $1/2$  it takes the process time at least a constant times  $x^2$  to reach  $B_r$ . Combining these two estimates will show that with probability of order  $1/x$  the walk takes time of order  $x^2$  to reach  $B_r$ , which gives the desired tail bound. Roughly speaking, the first estimate (reaching distance  $x$ ) is provided by the optional stopping theorem and the fact that  $\|X_k\|$  is a submartingale (cf. [7, Theorem 2.3]), and the second (taking quadratic time to return) is provided by a maximal inequality applied to an appropriate quadratic displacement functional (cf. [7, Lemma 4.11]). A technicality required for the first estimate is that to apply optional stopping, we need uniform integrability; so we actually work with a truncated version of  $\|X_k\|$ .

We now give the details. Recall that  $R_k = \|X_k\|$  and let  $\mathcal{F}_k = \sigma(X_0, \dots, X_k)$ . Lemmas 5.1 and 5.2, with the fact that  $V > U$  by (A4), imply that

$$\mathbb{E}[R_{k+1} - R_k \mid \mathcal{F}_k] \geq \frac{2\varepsilon}{R_k} + o(R_k^{-1}) \geq \frac{\varepsilon}{R_k}, \quad (6.6)$$

for all  $R_k > r_1$ , for sufficiently large  $r_1 \geq r_0$  and some positive constant  $\varepsilon$ . Now, suppose that  $r$  and  $R$  satisfy  $R > r > r_1$  and fix  $x$  with  $x \gg R$ . Set  $R_k^x := \min\{2x, R_k\}$  and  $\sigma_x := \min\{k \geq 0 : R_k > x\}$ . Since  $X_k$  is a martingale, we have that  $R_k$  is a submartingale, as is the stopped process  $Y_k := R_{k \wedge \tau_r \wedge \sigma_x}$ . In order to achieve uniform integrability, we consider the truncated process  $Y_k^x := R_{k \wedge \tau_r \wedge \sigma_x}^x$  and show that this is a submartingale.

For  $k \geq \tau_r \wedge \sigma_x$ , we have  $Y_{k+1}^x - Y_k^x = 0$  so  $\mathbb{E}[Y_{k+1}^x - Y_k^x \mid \mathcal{F}_k] = 0$ . For  $k < \tau_r \wedge \sigma_x$ ,

$$Y_{k+1}^x - Y_k^x = R_{k+1} - R_k + (2x - R_{k+1})\mathbf{1}\{R_{k+1} > 2x\},$$

and the last term can be bounded in absolute value:

$$\begin{aligned} |(2x - R_{k+1})\mathbf{1}\{R_{k+1} > 2x\}| &\leq |R_{k+1} - R_k|\mathbf{1}\{R_{k+1} > 2x\} \\ &\leq |R_{k+1} - R_k|\mathbf{1}\{|R_{k+1} - R_k| > x\} \\ &\leq |R_{k+1} - R_k|^p x^{1-p}, \end{aligned}$$

for  $p > 2$  as appearing in (A0), since on  $\{k < \sigma_x\}$  we have  $R_k < x$  and therefore  $R_{k+1} > 2x$  implies that  $|R_{k+1} - R_k| > x$ . Applying (5.2) from Lemma 5.1 we obtain

$$\mathbb{E}[|(Y_{k+1}^x - Y_k^x) - (R_{k+1} - R_k)| \mid \mathcal{F}_k] \leq Bx^{1-p},$$

for some  $B < \infty$  not depending on  $x$ . Combining this with (6.6) and again the fact that  $R_k < x$  on  $\{k < \sigma_x\}$ , we have that

$$\mathbb{E}[Y_{k+1}^x - Y_k^x \mid \mathcal{F}_k] \geq \frac{\varepsilon}{R_k} - Bx^{1-p} \geq \frac{\varepsilon}{x} - Bx^{1-p} \geq 0,$$

for sufficiently large  $x$ .

Hence, for sufficiently large  $x$ ,  $Y_k^x$  is a uniformly integrable submartingale and therefore, given  $X_0 \notin B_R$ , by optional stopping,

$$\begin{aligned} R < R_0 = Y_0^x &\leq \mathbb{E}[Y_{\sigma_x \wedge \tau_r}^x \mid X_0] = \mathbb{E}[Y_{\sigma_x}^x \mathbf{1}\{\sigma_x < \tau_r\} \mid X_0] + \mathbb{E}[Y_{\tau_r}^x \mathbf{1}\{\tau_r < \sigma_x\} \mid X_0] \\ &\leq 2x\mathbb{P}[\sigma_x < \tau_r \mid X_0] + r. \end{aligned}$$

In other words, given  $X_0 \notin B_R$ ,

$$\mathbb{P}\left[\max_{0 \leq k \leq \tau_r} R_k > x \mid X_0\right] \geq \frac{R-r}{2x}, \quad (6.7)$$

for all sufficiently large  $x$ .

Now, consider  $W_k := R_{\sigma_x+k} - R_{\sigma_x}$ , adapted to  $\mathcal{G}_k := \mathcal{F}_{\sigma_x+k}$ . We have

$$W_{k+1}^2 - W_k^2 = R_{\sigma_x+k+1}^2 - R_{\sigma_x+k}^2 - 2R_{\sigma_x}(R_{\sigma_x+k+1} - R_{\sigma_x+k}).$$

Using the fact that  $R_k$  is a submartingale together with the strong Markov property for  $X$  at the stopping time  $\sigma_x + k$  yields  $\mathbb{E}[R_{\sigma_x+k+1} - R_{\sigma_x+k} \mid \mathcal{F}_{\sigma_x+k}] \geq 0$  a.s., and Lemmas 5.1 and 5.2 again with the strong Markov property imply that  $\mathbb{E}[R_{\sigma_x+k+1}^2 - R_{\sigma_x+k}^2 \mid \mathcal{F}_{\sigma_x+k}] \leq C$  a.s., for some constant  $C < \infty$ ; hence  $\mathbb{E}[W_{k+1}^2 - W_k^2 \mid \mathcal{G}_k] \leq C$  a.s., for some constant  $C < \infty$ . Then a maximal inequality [13, Lemma 3.1] similar to Doob's submartingale inequality implies that, on  $\{\sigma_x < \infty\}$ ,

$$\mathbb{P}\left[\max_{0 \leq k \leq n} W_k^2 \geq y \mid \mathcal{G}_0\right] \leq \frac{Cn}{y}, \text{ for any } y > 0.$$

In particular, we may choose  $\varepsilon > 0$  small enough so that

$$\mathbb{P}\left[\max_{0 \leq k \leq \varepsilon x^2} |R_{\sigma_x+k} - R_{\sigma_x}| \geq x/2 \mid \mathcal{F}_{\sigma_x}\right] \leq \frac{1}{2}, \text{ on } \{\sigma_x < \infty\}. \quad (6.8)$$

Combining the inequalities (6.7) and (6.8), we find that given  $X_0 \notin B_R$ ,

$$\begin{aligned} &\mathbb{P}\left[\left\{\max_{0 \leq k \leq \tau_r} R_k > x\right\} \cap \left\{\max_{0 \leq k \leq \varepsilon x^2} |R_{\sigma_x+k} - R_{\sigma_x}| < x/2\right\} \mid X_0\right] \\ &= \mathbb{E}\left[\mathbf{1}\{\sigma_x < \tau_r\} \mathbb{P}\left[\max_{0 \leq k \leq \varepsilon x^2} |R_{\sigma_x+k} - R_{\sigma_x}| < x/2 \mid \mathcal{F}_{\sigma_x}\right] \mid X_0\right] \\ &\geq \frac{1}{2} \mathbb{P}\left[\max_{0 \leq k \leq \tau_r} R_k > x \mid X_0\right] \geq \frac{R-r}{4x}, \end{aligned}$$

for sufficiently large  $x$ , where the equality here uses the fact that  $\{\sigma_x < \tau_r\} \in \mathcal{F}_{\sigma_x}$ .

If both of the events  $\{\max_{0 \leq k \leq \tau_r} R_k > x\}$  and  $\{\max_{0 \leq k \leq \varepsilon x^2} |R_{\sigma_x+k} - R_{\sigma_x}| < x/2\}$  occur, then the process  $X_k$  leaves the ball  $B_x$  before time  $\tau_r$  and takes more than  $\varepsilon x^2$  steps to return to the ball  $B_{x/2} \subset B_r$ , and therefore  $\tau_r > \varepsilon x^2$ . Setting  $m = \varepsilon x^2$  and  $c = (R-r)\sqrt{\varepsilon}/4$  yields the claimed inequality.  $\square$

*Remark 6.3.* It is only in the proof of Lemma 6.2 that we use the condition  $U < V$  from (A4). In the case  $U = V$ , inequality (6.6) holds only for (any)  $\varepsilon < 0$ , and not  $\varepsilon > 0$ ; thus to obtain a submartingale one should look at  $(Y_k^x)^\gamma$  for  $\gamma > 1$ . The modified argument yields a weaker version of (6.1), with  $m^{-1/2}$  replaced by  $m^{-(1/2)-\delta}$  for any  $\delta > 0$ , but, as stated in Remark 2.5, this is still comfortably enough to give Theorem 2.4 (any exponent greater than  $-1$  in the tail bound will do). We omit these additional technical details, as the case  $U = V$  is outside our main interest.

## A Recurrence in one dimension

We use a Lyapunov function method with function  $f(x) = \log(1 + |x|)$ .

**Lemma A.1.** *Suppose that  $X$  is a discrete-time, time-homogeneous Markov process on  $\mathbb{X} \subseteq \mathbb{R}$ . Suppose that for some  $p > 2$  and  $v > 0$ ,*

$$\begin{aligned} \sup_{x \in \mathbb{X}} \mathbb{E}[(X_{n+1} - X_n)^p \mid X_n = x] &< \infty; \\ \inf_{x \in \mathbb{X}} \mathbb{E}[(X_{n+1} - X_n)^2 \mid X_n = x] &\geq v. \end{aligned}$$

*Suppose also that for some bounded set  $A \subset \mathbb{R}$ ,*

$$\mathbb{E}[X_{n+1} - X_n \mid X_n = x] = 0, \text{ for all } x \in \mathbb{X} \setminus A.$$

*Then there exists a bounded set  $A' \subset \mathbb{R}$  for which*

$$\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq 0, \text{ for all } x \in \mathbb{X} \setminus A'.$$

*Proof.* Write  $\Delta = X_1 - X_0$  and  $E_x = \{|\Delta| < |x|\}$ . We compute

$$\begin{aligned} \mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] &= \mathbb{E}_x[(f(x + \Delta) - f(x))\mathbf{1}(E_x)] \\ &\quad + \mathbb{E}_x[(f(x + \Delta) - f(x))\mathbf{1}(E_x^c)]. \end{aligned}$$

On  $\{|\Delta| < |x|\}$  we have that  $x$  and  $x + \Delta$  have the same sign, so

$$\begin{aligned} &\mathbb{E}_x[(f(x + \Delta) - f(x))\mathbf{1}(E_x)] \\ &= \mathbb{E}_x \left[ \log \left( \frac{1 + |x + \Delta|}{1 + |x|} \right) \mathbf{1}(E_x) \right] \\ &= \mathbb{E}_x \left[ \log \left( 1 + \frac{\Delta \operatorname{sgn}(x)}{1 + |x|} \right) \mathbf{1}(E_x) \right] \\ &\leq \left( \frac{\operatorname{sgn}(x)}{1 + |x|} \right) \mathbb{E}_x[\Delta \mathbf{1}(E_x)] - \frac{1}{6}(1 + |x|)^{-2} \mathbb{E}_x[\Delta^2 \mathbf{1}(E_x)], \end{aligned}$$

using the inequality  $\log(1 + y) \leq y - \frac{1}{6}y^2$  for all  $-1 < y \leq 1$ . Here, since  $\mathbb{E}_x[\Delta] = 0$  for  $x \notin A$ ,

$$|\mathbb{E}_x[\Delta \mathbf{1}(E_x)]| \leq \mathbb{E}_x[|\Delta| \mathbf{1}(E_x^c)] \leq \mathbb{E}_x[|\Delta|^p |x|^{1-p}] = o(|x|^{-1}).$$

Similarly,

$$\mathbb{E}_x[\Delta^2 \mathbf{1}(E_x)] \geq v - \mathbb{E}_x[\Delta^2 \mathbf{1}(E_x^c)] \geq v - o(1).$$

(Note that here, and in what follows, our notation follows the convention as described by (5.1); consequently, in one dimension the error terms are understood to be uniform as either  $x \rightarrow +\infty$ , or  $x \rightarrow -\infty$ .) Finally we estimate the term

$$|\mathbb{E}_x[(f(x + \Delta) - f(x))\mathbf{1}(E_x^c)]| \leq \mathbb{E}_x[(\log(1 + |\Delta|) + \log(1 + 2|\Delta|))\mathbf{1}(E_x^c)].$$

Here,

$$\begin{aligned} \log(1 + 2|\Delta|)\mathbf{1}(E_x^c) &= \log(1 + 2|\Delta|)|\Delta|^p|\Delta|^{-p}\mathbf{1}(E_x^c) \\ &\leq |x|^{-p}\log(1 + 2|x|)|\Delta|^p, \end{aligned}$$

for all  $x$  with  $|x|$  greater than some  $x_0$  sufficiently large, using the fact that  $y \mapsto y^{-p}\log(1 + 2y)$  is eventually decreasing. It follows that

$$|\mathbb{E}_x[(f(x + \Delta) - f(x))\mathbf{1}(E_x^c)]| \leq 2|x|^{-p}\log(1 + 2|x|)\mathbb{E}_x[|\Delta|^p] = o(|x|^{-2}).$$

Combining these calculations we obtain

$$\begin{aligned} \mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] &\leq \left( \frac{\operatorname{sgn}(x)}{1 + |x|} \right) o(|x|^{-1}) - \frac{1}{6}(1 + |x|)^{-2}(v - o(1)) + o(|x|^{-2}) \\ &\leq -\frac{v}{6}(1 + |x|)^{-2} + o(|x|^{-2}), \end{aligned}$$

which is negative for all  $x$  with  $|x|$  sufficiently large.  $\square$

*Proof of Theorem 2.2.* Under assumptions (A0), (A1) and (A2), the hypotheses of Lemma A.1 are satisfied, so that for some  $x_0 \in \mathbb{R}_+$ ,

$$\mathbb{E}[f(X_{n+1}) - f(X_n) \mid X_n = x] \leq 0, \text{ for all } x \in \mathbb{X} \text{ with } |x| \geq x_0,$$

where  $f(x) = \log(1 + |x|)$ .

We note that assumption (A0) implies that  $\mathbb{E}[|X_n|] < \infty$  for all  $n$ , and therefore  $\mathbb{E}[f(X_n)] < \infty$  for all  $n$ . Let  $n_0 \in \mathbb{N}$  and set  $\tau = \min\{n \geq n_0 : |X_n| \leq x_0\}$ . Let  $Y_n = f(X_{n \wedge \tau})$ . Then  $(Y_n, n \geq n_0)$  is a non-negative supermartingale, and hence there exists a random variable  $Y_\infty \in \mathbb{R}_+$  with  $\lim_{n \rightarrow \infty} Y_n = Y_\infty$ , a.s. In particular, this means that

$$\limsup_{n \rightarrow \infty} f(X_n) \leq Y_\infty, \text{ on } \{\tau = \infty\}.$$

Setting  $\zeta = \sup\{|x| : x \in \mathbb{X}, f(x) \leq Y_\infty\}$ , which satisfies  $\zeta < \infty$ , a.s., since  $f(x) \rightarrow \infty$  as  $|x| \rightarrow \infty$ , it follows that  $\limsup_{n \rightarrow \infty} |X_n| \leq \zeta$  on  $\{\tau = \infty\}$ . However, under assumptions (A0), (A1) and (A2), Proposition 2.1 implies that  $\limsup_{n \rightarrow \infty} |X_n| = +\infty$ , a.s., so to avoid contradiction, we must have  $\tau < \infty$ , a.s. In other words,

$$\mathbb{P}\left[\inf_{n \geq n_0} |X_n| \leq x_0\right] = 1,$$

and since  $n_0$  was arbitrary, it follows that

$$\mathbb{P}\left[\bigcap_{n_0 \in \mathbb{N}} \left\{ \inf_{n \geq n_0} |X_n| \leq x_0 \right\}\right] = 1,$$

which gives the result.  $\square$

## Acknowledgements

Part of this work was supported by the Engineering and Physical Sciences Research Council [grant number EP/J021784/1].

An antecedent of this work, concerning only the elliptic random walk in two dimensions, was written down in 2008–9 by MM and AW, who benefited from stimulating discussions with Iain MacPhee (7/11/1957–13/1/2012). The present authors also thank Stas Volkov for a comment that inspired Remark 3.3(d).

## References

- [1] H.C. Berg, *Random Walks in Biology*. Expanded Edition, Princeton University Press, Princeton, 1993.
- [2] N.H. Bingham, Random walk on spheres. *Z. Wahrschein. verw. Gebiete* **22** (1972), 169–192.
- [3] B. Carazza, The history of the random-walk problem: considerations on the interdisciplinarity in modern physics. *Rivista del Nuovo Cimento Serie 2* **7** (1977) 419–427.
- [4] K.L. Chung and W.H.J. Fuchs, On the distribution of values of sums of random variables. *Mem. Amer. Math. Soc.* **6** (1951) 12pp.
- [5] A. Gut, *Probability: A Graduate Course*, Springer, Uppsala, 2005.
- [6] O. Hryniv, I.M. MacPhee, M.V. Menshikov, and A.R. Wade, Non-homogeneous random walks with non-integrable increments and heavy-tailed random walks on strips. *Electr. J. Probab.* **17** (2012) article 59, 28pp.
- [7] O. Hryniv, M.V. Menshikov, and A.R. Wade, Excursions and path functionals for stochastic processes with asymptotically zero drifts. *Stoch. Process. Appl.* **123** (2013) 1891–1921.
- [8] B.D. Hughes, *Random Walks and Random Environments. Volume 1: Random Walks*. Clarendon Press, Oxford, 1995.
- [9] O. Kallenberg, *Foundations of Modern Probability*. 2nd ed., Springer, New York, 2002.
- [10] J.H.B. Kemperman, The oscillating random walk. *Stoch. Process. Appl.* **2** (1974) 1–29.
- [11] J.F.C. Kingman, Random walks with spherical symmetry. *Acta Math.* **109** (1963) 11–63.
- [12] J. Lamperti, Criteria for the recurrence and transience of stochastic processes I. *J. Math. Anal. Appl.* **1** (1960) 314–330.
- [13] M.V. Menshikov, M. Vachkovskaia, and A.R. Wade, Asymptotic behaviour of randomly reflecting billiards in unbounded tubular domains. *J. Statist. Phys.* **132** (2008) 1097–1133.

- [14] R. Nossal, Stochastic aspects of biological locomotion. *J. Statist. Phys.* **30** (1983) 391–400.
- [15] R.J. Nossal and G.H. Weiss, A generalized Pearson random walk allowing for bias. *J. Statist. Phys.* **10** (1974) 245–253.
- [16] K. Pearson and Lord Rayleigh, The problem of the random walk. *Nature* **72** (1905) pp. 294, 318, 342.
- [17] K. Pearson, *A Mathematical Theory of Random Migration*. Drapers' Company Research Memoirs, Dulau and Co., London, 1906.
- [18] Y. Peres, S. Popov, and P. Sousi, On recurrence and transience of self-interacting random walks. *Bull. Brazilian Math. Soc.* **44** (2013) 841–867.
- [19] B.A. Rogozin and S.G. Foss, The recurrence of an oscillating random walk. *Theor. Probability Appl.* **23** (1978) 155–162. Translated from *Teor. Veroyatn. Primen.* **23** (1978) 161–169.
- [20] O. Zeitouni, *Lecture Notes on Random Walks in Random Environment*. 9th May 2006 version, <http://www.wisdom.weizmann.ac.il/~zeitouni/>.